



0590
0514

RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/992,095A
Source: OIPB
Date Processed by STIC: 5/13/02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to :
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
Box Sequence, Room 1B03 Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202



Does Not Comply
Corrected Diskette Needed

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/992,095A

DATE: 05/13/2002
TIME: 16:15:14

See pp 7 & 6

Input Set : F:\Seqlist.txt
Output Set: N:\CRF3\05132002\I992095A.raw

3 <110> APPLICANT: Benjanin, Stephane
4 Tanaka, Hiroaki
6 <120> TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
8 <130> FILE REFERENCE: 91.US5.DIV
10 <140> CURRENT APPLICATION NUMBER: US 09/992,095A
C--> 11 <141> CURRENT FILING DATE: 2002-05-01
13 <150> PRIOR APPLICATION NUMBER: US 09/924,340
14 <151> PRIOR FILING DATE: 2001-08-06
16 <150> PRIOR APPLICATION NUMBER: PCT/IB01/01715
17 <151> PRIOR FILING DATE: 2001-08-06
19 <150> PRIOR APPLICATION NUMBER: US 60/305,456
20 <151> PRIOR FILING DATE: 2001-07-13
22 <150> PRIOR APPLICATION NUMBER: US 60/302,277
23 <151> PRIOR FILING DATE: 2001-06-29
25 <160> NUMBER OF SEQ ID NOS: 112
27 <170> SOFTWARE: JPatent
29 <210> SEQ ID NO: 1
30 <211> LENGTH: 2016
31 <212> TYPE: DNA
32 <213> ORGANISM: Homo sapiens
34 <220> FEATURE:
35 <221> NAME/KEY: 5'UTR
36 <222> LOCATION: 1..1434
38 <220> FEATURE:
39 <221> NAME/KEY: CDS
40 <222> LOCATION: 1435..1836
42 <220> FEATURE:
43 <221> NAME/KEY: 3'UTR
44 <222> LOCATION: 1837..2016
46 <220> FEATURE:
47 <221> NAME/KEY: polyA_signal
48 <222> LOCATION: 1965..1970
50 <220> FEATURE:
51 <221> NAME/KEY: polyA_site
52 <222> LOCATION: 2001..2016
54 <400> SEQUENCE: 1
55 aaggtctctc tgcattgcata caccaaggaa aagccacatg aggacataac caggaagaga 60
56 gccatcacca agaaccogaa catgocggaca ccctgatctc ggacttctag ccttcagaac 120
57 cgttgccaca gttttgatga tcatctctct cccaaccaag atggtggaaa aagcaaaaac 180
58 gtggtgaatc ttggagcaat cggacaaggc atgaaacgct tccaatttct gttaaactgc 240
59 tgtgagccag ggacaattcc tgaatgcctcc atcctagcag ctgccttgga tctactatgc 300
60 ggcattcttc tgattcattt ttctccattt gtgctgtttt tctctgtgat gtgaatccat 360
61 ccctatccat tatgtcatgc ctccatcttt tgctgcttct tcagattgca ctgagccata 420

RAW SEQUENCE LISTING

DATE: 05/13/2002

PATENT APPLICATION: US/09/992,095A

TIME: 16:15:14

Input Set : F:\Seqlist.txt

Output Set: N:\CRF3\05132002\I992095A.raw

```

62 agaggaagcc cctgtggtgg ccagagcagc cttgttctctg gaatgtgctc gttttgttca 480
63 ccgctgcaac cgtggcaact ggccagagtg gatgaaaggg caccacgtga acatcaccaa 540
64 gaaaggactt tcccggggac gctctcccat tgtgggcaac aagcgaaacc agaagctgca 600
65 gtggaatgca gccaagctct tctaccaatg gggagacaag gaaaaaaggt gaagaataaa 660
66 aggaaattca agaggaccaa gttttctgcta atttttagaca gagctgaaca taaacacaca 720
67 taaagagggtt ccatatatct ctcttttctt aaagattact tggaataact gttacaattt 780
68 ccgttaataa ttcagctgaa tgtgtctacc aatgtgctta ccaactaagg caattggcgt 840
69 ccgattgaat gagctgtgcc acggggaaag tgagagccca gccaacctgc tgggtctcat 900
70 ttacgatgaa gagaccaaga ggagacttag aaaggaggat gaggaggaag acttttttaga 960
71 tgacattcca ctttcaagtc aatacacagc tcatcttgca tttaaaagct gattatggtg 1020
72 caagcaactt tcgggctgga aattctacag aagcttgtct tttccattct tgatgagagg 1080
73 caaagtcccc ggcaacaaat taactcagga gagaaaatgg ttttctgaa aaaaacgata 1140
74 gcttaaatat ctacagaaag accgtaattt ccacctattt tcaaataaaa tcgtgaaaaa 1200
75 cacatttgga ctagagctga aacaacttca ctgccctcaa aacagcaaga cagacatccc 1260
76 tcataaaatg aactgacaga atttttatag ctccaaatct agttcactgc catatacata 1320
77 gtctaaatct gattgaatag cagcgtagaa atcttgcgaa attacttccc atttctgttt 1380
78 tcgttaaaag gtactgtgaa cccctctaaa tgcggttgcc ctttgcctt gaag atg 1437
79 Met
80 1
81 gca gca tgt cag ctt ctt ctg gag att acc acc ttc ctg cga gag acc 1485
82 Ala Ala Cys Gln Leu Leu Leu Glu Ile Thr Thr Phe Leu Arg Glu Thr
83 5 10 15
84 ttt tct tgc ctg ccc aga cct cgc act gag cct ctg gtg gct tca acg 1533
85 Phe Ser Cys Leu Pro Arg Pro Arg Thr Glu Pro Leu Val Ala Ser Thr
86 20 25 30
87 gac cac acc aaa atg cca tct caa atg gaa cac gcc atg gaa acc atg 1581
88 Asp His Thr Lys Met Pro Ser Gln Met Glu His Ala Met Glu Thr Met
89 35 40 45
90 atg ttt aca ttt cac aaa ttc gct ggg gat aaa ggc tac tta aca aag 1629
91 Met Phe Thr Phe His Lys Phe Ala Gly Asp Lys Gly Tyr Leu Thr Lys
92 50 55 60 65
93 gag gac ctg aga gta ctc atg gaa aag gag ttc cct gga ttt ttg gaa 1677
94 Glu Asp Leu Arg Val Leu Met Glu Lys Glu Phe Pro Gly Phe Leu Glu
95 70 75 80
96 aat caa aaa gac cct ctg gct gtg gac aaa ata atg aag gac ctg gac 1725
97 Asn Gln Lys Asp Pro Leu Ala Val Asp Lys Ile Met Lys Asp Leu Asp
98 85 90 95
99 cag tgt aga gat ggc aaa gtg ggc ttc cag agc ttc ttt tcc cta att 1773
100 Gln Cys Arg Asp Gly Lys Val Gly Phe Gln Ser Phe Phe Ser Leu Ile
101 100 105 110
102 gcg ggc ctc acc att gca tgc aat gac tat ttt gta gta cac atg aag 1821
103 Ala Gly Leu Thr Ile Ala Cys Asn Asp Tyr Phe Val Val His Met Lys
104 115 120 125
105 cag aag gga aag aag taggcagaaa tgagcagttc gtcctccct gataagagtt 1876
106 Gln Lys Gly Lys Lys
107 130
108 gtcccaaagg gtcgcttaag gaatctgccc cacagcttcc cccatagaag gatttcatga 1936
109 gcagatcagg acacttagca aatgtaaaaa taaaatctaa ctctcatttg agaagcagag 1996

```

RAW SEQUENCE LISTING

DATE: 05/13/2002

PATENT APPLICATION: US/09/992,095A

TIME: 16:15:14

Input Set : F:\Seqlist.txt

Output Set: N:\CRF3\05132002\I992095A.raw

```

112 <210> SEQ ID NO: 2
113 <211> LENGTH: 134
114 <212> TYPE: PRT
115 <213> ORGANISM: Homo sapiens
117 <400> SEQUENCE: 2
118 Met Ala Ala Cys Gln Leu Leu Leu Glu Ile Thr Thr Phe Leu Arg Glu
119 1 5 10 15
120 Thr Phe Ser Cys Leu Pro Arg Pro Arg Thr Glu Pro Leu Val Ala Ser
121 20 25 30
122 Thr Asp His Thr Lys Met Pro Ser Gln Met Glu His Ala Met Glu Thr
123 35 40 45
124 Met Met Phe Thr Phe His Lys Phe Ala Gly Asp Lys Gly Tyr Leu Thr
125 50 55 60
126 Lys Glu Asp Leu Arg Val Leu Met Glu Lys Glu Phe Pro Gly Phe Leu
127 65 70 75 80
128 Glu Asn Gln Lys Asp Pro Leu Ala Val Asp Lys Ile Met Lys Asp Leu
129 85 90 95
130 Asp Gln Cys Arg Asp Gly Lys Val Gly Phe Gln Ser Phe Phe Ser Leu
131 100 105 110
132 Ile Ala Gly Leu Thr Ile Ala Cys Asn Asp Tyr Phe Val Val His Met
133 115 120 125
134 Lys Gln Lys Gly Lys Lys
135 130
137 <210> SEQ ID NO: 3
138 <211> LENGTH: 1081
139 <212> TYPE: DNA
140 <213> ORGANISM: Homo sapiens
142 <220> FEATURE:
143 <221> NAME/KEY: 5'UTR
144 <222> LOCATION: 1..38
146 <220> FEATURE:
147 <221> NAME/KEY: CDS
148 <222> LOCATION: 39..917
150 <220> FEATURE:
151 <221> NAME/KEY: 3'UTR
152 <222> LOCATION: 918..1081
154 <220> FEATURE:
155 <221> NAME/KEY: polyA_signal
156 <222> LOCATION: 1045..1050
158 <220> FEATURE:
159 <221> NAME/KEY: polyA_site
160 <222> LOCATION: 1066..1081
162 <400> SEQUENCE: 3
163 gtccagcctg ttgctgatgc tgccgtgcgg tacttgtc atg gag ctg gca ctg cgg 56
164 Met Glu Leu Ala Leu Arg
165 -25 -20
166 cgc tct ccc gtc ccg cgg tgg ttg ctg ctg ccg ctg ctg ctg ggc 104
167 Arg Ser Pro Val Pro Arg Trp Leu Leu Leu Leu Pro Leu Leu Gly
168 -15 -10 -5

```

DATE: 05/13/2002

TIME: 16:15:14

Input Set : F:\Seqlist.txt

Output Set: N:\CRF3\05132002\I992095A.raw

169	ctg	aac	gca	gga	gct	gtc	att	gac	tgg	ccc	aca	gag	gag	ggc	aag	gaa	152
170	Leu	Asn	Ala	Gly	Ala	Val	Ile	Asp	Trp	Pro	Thr	Glu	Glu	Gly	Lys	Glu	
171			1					5					10				
172	gta	tgg	gat	tat	gtg	acg	gtc	cgc	aag	gat	gcc	tac	atg	ttc	tgg	tgg	200
173	Val	Trp	Asp	Tyr	Val	Thr	Val	Arg	Lys	Asp	Ala	Tyr	Met	Phe	Trp	Trp	
174		15					20					25					
175	ctc	tat	tat	gcc	acc	aac	tcc	tgc	aag	aac	ttc	tca	gaa	ctg	ccc	ctg	248
176	Leu	Tyr	Tyr	Ala	Thr	Asn	Ser	Cys	Lys	Asn	Phe	Ser	Glu	Leu	Pro	Leu	
177	30					35					40					45	
178	gtc	atg	tgg	ctt	cag	ggc	ggg	cca	ggc	ggg	tct	agc	act	gga	ttt	gga	296
179	Val	Met	Trp	Leu	Gln	Gly	Gly	Pro	Gly	Gly	Ser	Ser	Thr	Gly	Phe	Gly	
180					50					55					60		
181	aac	ttt	gag	gaa	att	ggg	ccc	ctt	gac	agt	gat	ctc	aaa	cca	cgg	aaa	344
182	Asn	Phe	Glu	Glu	Ile	Gly	Pro	Leu	Asp	Ser	Asp	Leu	Lys	Pro	Arg	Lys	
183				65					70					75			
184	acc	acc	tgg	ctc	cag	gct	gcc	agt	ctc	cta	ttt	gtg	gat	aat	ccc	gtg	392
185	Thr	Thr	Trp	Leu	Gln	Ala	Ala	Ser	Leu	Leu	Phe	Val	Asp	Asn	Pro	Val	
186			80					85					90				
187	ggc	act	ggg	ttc	agt	tat	gtg	aat	ggg	agt	ggg	gcc	tat	gcc	aag	gac	440
188	Gly	Thr	Gly	Phe	Ser	Tyr	Val	Asn	Gly	Ser	Gly	Ala	Tyr	Ala	Lys	Asp	
189		95					100					105					
190	ctg	gct	atg	gtg	gct	tca	gac	atg	atg	gtt	ctc	ctg	aag	acc	ttc	ttc	488
191	Leu	Ala	Met	Val	Ala	Ser	Asp	Met	Met	Val	Leu	Leu	Lys	Thr	Phe	Phe	
192	110					115					120					125	
193	agt	tgc	cac	aaa	gaa	ttc	cag	aca	gtt	cca	ttc	tac	att	ttc	tca	gag	536
194	Ser	Cys	His	Lys	Glu	Phe	Gln	Thr	Val	Pro	Phe	Tyr	Ile	Phe	Ser	Glu	
195				130						135					140		
196	tcc	tat	gga	gga	aaa	atg	gca	gct	ggc	att	ggg	cta	gag	ctt	tat	aag	584
197	Ser	Tyr	Gly	Gly	Lys	Met	Ala	Ala	Gly	Ile	Gly	Leu	Glu	Leu	Tyr	Lys	
198			145						150					155			
199	gcc	att	cag	cga	ggg	acc	atc	aag	tgc	aac	ttt	gcg	ggg	gtt	gcc	ttg	632
200	Ala	Ile	Gln	Arg	Gly	Thr	Ile	Lys	Cys	Asn	Phe	Ala	Gly	Val	Ala	Leu	
201			160					165					170				
202	ggg	gat	tcc	tgg	atc	tcc	cct	gtt	gat	tcg	gtg	ctc	tcc	tgg	gga	cct	680
203	Gly	Asp	Ser	Trp	Ile	Ser	Pro	Val	Asp	Ser	Val	Ser	Trp	Gly	Pro		
204		175					180					185					
205	tac	ctg	tac	agc	atg	tct	ctt	ctc	gaa	gac	aaa	ggg	ctg	gca	gag	gtg	

RAW SEQUENCE LISTING

DATE: 05/13/2002

PATENT APPLICATION: US/09/992,095A

TIME: 16:15:14

Input Set : F:\Seqlist.txt

Output Set: N:\CRF3\05132002\I992095A.raw

```

218 Gly Tyr Arg Ala His Gly Trp Cys Cys Gln Thr Trp Ser Leu His
219      255                      260                      265
220 tgaggctccc cacatatctg caaatgattg catgctggat aataaatctc ttgggtotaa 977
221 gcagtgatgt agtggctcct tacagagtca gaaagccacc caggcctgca agacttgctt 1037
222 gtccttcact aaatgtatgg attctattaa aaaaaaaaaa aaaa 1081
224 <210> SEQ ID NO: 4
225 <211> LENGTH: 293
226 <212> TYPE: PRT
227 <213> ORGANISM: Homo sapiens
229 <220> FEATURE:
230 <221> NAME/KEY: SIGNAL
231 <222> LOCATION: 1..26
233 <400> SEQUENCE: 4
234 Met Glu Leu Ala Leu Arg Arg Ser Pro Val Pro Arg Trp Leu Leu Leu
235      -25                      -20                      -15
236 Leu Pro Leu Leu Leu Gly Leu Asn Ala Gly Ala Val Ile Asp Trp Pro
237      -10                      -5                      1                      5
238 Thr Glu Glu Gly Lys Glu Val Trp Asp Tyr Val Thr Val Arg Lys Asp
239                      10                      15                      20
240 Ala Tyr Met Phe Trp Trp Leu Tyr Tyr Ala Thr Asn Ser Cys Lys Asn
241                      25                      30                      35
242 Phe Ser Glu Leu Pro Leu Val Met Trp Leu Gln Gly Gly Pro Gly Gly
243      40                      45                      50
244 Ser Ser Thr Gly Phe Gly Asn Phe Glu Glu Ile Gly Pro Leu Asp Ser
245      55                      60                      65                      70
246 Asp Leu Lys Pro Arg Lys Thr Thr Trp Leu Gln Ala Ala Ser Leu Leu
247                      75                      80                      85
248 Phe Val Asp Asn Pro Val Gly Thr Gly Phe Ser Tyr Val Asn Gly Ser
249                      90                      95                      100
250 Gly Ala Tyr Ala Lys Asp Leu Ala Met Val Ala Ser Asp Met Met Val
251                      105                      110                      115
252 Leu Leu Lys Thr Phe Phe Ser Cys His Lys Glu Phe Gln Thr Val Pro
253      120                      125                      130
254 Phe Tyr Ile Phe Ser Glu Ser Tyr Gly Gly Lys Met Ala Ala Gly Ile
255      135                      140                      145                      150
256 Gly Leu Glu Leu Tyr Lys Ala Ile Gln Arg Gly Thr Ile Lys Cys Asn
257                      155                      160                      165
258 Phe Ala Gly Val Ala Leu Gly Asp Ser Trp Ile Ser Pro Val Asp Ser
259                      170                      175                      180
260 Val Leu Ser Trp Gly Pro Tyr Leu Tyr Ser Met Ser Leu Leu Glu Asp
261                      185                      190                      195
262 Lys Gly Leu Ala Glu Val Ser Lys Val Ala Glu Gln Val Leu Asn Ala
263      200                      205                      210
264 Val Asn Lys Gly Leu Tyr Arg Glu Ala Thr Glu Leu Trp Gly Lys Ala
265      215                      220                      225                      230
266 Glu Met Ile Ile Glu Gln Val Lys Arg Gly Asn Thr Gln Arg Leu Ala
267                      235                      240                      245
268 Cys Leu Ala Phe Ser Gly Gly Tyr Arg Ala His Gly Trp Cys Cys Gln
269                      250                      255                      260

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/992,095A

DATE: 05/13/2002
TIME: 16:15:15

Input Set : F:\Seqlist.txt
Output Set: N:\CRF3\05132002\I992095A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:27; Xaa Pos. 116,233
Seq#:28; Xaa Pos. 116,233
Seq#:29; Xaa Pos. 30
Seq#:30; Xaa Pos. 29
Seq#:71; Xaa Pos. 157
Seq#:72; Xaa Pos. 156

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/992,095A

DATE: 05/13/2002
TIME: 16:15:15

✱

Input Set : F:\Seqlist.txt
Output Set: N:\CRF3\05132002\I992095A.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1471 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:27 ←
L:1471 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:502
L:1492 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:27 ←
L:1492 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:838
L:1538 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:112
L:1552 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:224
L:1593 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:29 ←
L:1593 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:500
L:1656 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:48
L:4090 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:71 ←
L:4090 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71 after pos.:529
L:4157 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72 after pos.:160